BISS Biodiversity Information Science and Standards

Biodiversity Information Science and Standards 4: e59119
doi: 10.3897/biss.4.59119

OPEN ACCESS

Conference Abstract

# Multi-domain Collection Management Simplified — the Finnish National Collection Management System Kotka

Mikko Heikkinen‡, Anniina Kuusijärvi‡, Ville-Matti Riihikoski‡, Leif Schulman‡

‡ Finnish Museum of Natural History LUOMUS, Helsinki, Finland

Corresponding author: Mikko Heikkinen (mikko.heikkinen@helsinki.fi), Anniina Kuusijärvi (anniina.kuusijarvi@helsinki.fi)

Received: 29 Sep 2020 | Published: 09 Oct 2020

Citation: Heikkinen M, Kuusijärvi A, Riihikoski V-M, Schulman L (2020) Multi-domain Collection Management Simplified — the Finnish National Collection Management System Kotka. Biodiversity Information Science and Standards 4: e59119. https://doi.org/10.3897/biss.4.59119

## Abstract

Many natural history museums share a common problem: a multitude of legacy collection management systems (CMS) and the difficulty of finding a new system to replace them. Kotka is a CMS developed starting in 2011 at the Finnish Museum of Natural History (Luomus) and Finnish Biodiversity Information Facility (FinBIF) (Heikkinen et al. 2019, Schulman et al. 2019) to solve this problem. It has grown into a national system used by all natural history museums in Finland, and currently contains over two million specimens from several domains (zoological, botanical, paleontological, microbial, tissue sample and botanic garden collections).

Kotka is a web application where data can be entered, edited, searched and exported through a browser-based user interface. It supports designing and printing specimen labels, handling collection metadata and specimen transactions, and helps support Nagoya protocol compliance.

Creating a shared system for multiple institutions and collection types is difficult due to differences in their current processes, data formats, future needs and opinions. The more independent actors there are involved, the more complicated the development becomes. Successful development requires some trade-offs. Kotka has chosen features and

© Heikkinen M et al. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

development principles that emphasize fast development into a multitude of different purposes.

Kotka was developed using agile methods with a single person (a product owner) making development decisions, based on e.g., strategic objectives, customer value and user feedback. Technical design emphasizes efficient development and usage over completeness and formal structure of the data. It applies simple and pragmatic approaches and improves collection management by providing practical tools for the users. In these regards, Kotka differs in many ways from a traditional CMS.

Kotka stores data in a mostly denormalized free text format and uses a simple hierarchical data model. This allows greater flexibility and makes it easy to add new data fields and structures based on user feedback. Data harmonization and quality assurance is a continuous process, instead of doing it before entering data into the system. For example, specimen data with a taxon name can be entered into Kotka before the taxon name has been entered into the accompanying FinBIF taxonomy database.

*Example: simplified data about two specimens in Kotka, which have not been fully harmonized yet.*

- Taxon: *Corvus corone cornix*
- Country: FI
- Collector: Doe, John
- Coordinates: 668, 338
- Coordinate system: Finnish uniform coordinate system

- Taxon: *Corvus cornix*
- Country: Finland
- Collector: Doe, J.
- Coordinates: 60.2442, 25,7201
- Coordinate system: WGS84

Kotka's data model does not follow standards, but has grown organically to reflect practical needs from the users. This is true particularly of data collected in research projects, which are often unique and complicated (e.g. complex relationships between species), requiring new data fields and/or storing data as free text. The majority of the data can be converted into simplified standard formats (e.g. Darwin Core) for sharing. The main challenge with this has been vague definitions of many data sharing formats (e.g. Darwin Core, CETAF Specimen Preview Profile (CETAF 2020), allowing different interpretations.

Kotka trusts its users: it places very few limitations on what users can do, and has very simple user role management. Kotka stores the full history of all data, which allows fixing any possible errors and prevents data loss.

Kotka is open source software, but is tightly coupled with the infrastructure of the Finnish Biodiversity Information Facility (FinBIF). Currently, it is only offered as an online service (Software as a Service) hosted by FinBIF. However, it could be developed into a more

modular system that could, for example, utilize multiple different database backends and taxonomy data sources.

## Keywords

collection management systems, software, biodiversity informatics, natural history museums, botanic gardens

## Presenting author

Mikko Heikkinen

## Presented at

TDWG 2020

## References


- CETAF (2020) CETAF Specimen Preview Profile (CSPP). https://cetafidentifiers.biowikifarm.net/wiki/CETAF_Specimen_Preview_Profile_(CSPP). Accessed on: 2020-9-10.
- Heikkinen M, Riihikoski V, Kuusijrvi A, Talvitie D, Lahti T, Schulman L (2019) Kotka - A national multi-purpose collection management system. Biodiversity Information Science and Standards 3 https://doi.org/10.3897/biss.3.37179
- Schulman L, Juslén A, Lahti K (2019) FinBIF: An all-embracing, integrated, cross-sectoral biodiversity data infrastructure. Biodiversity Information Science and Standards 3: e37253. https://doi.org/10.3897/biss.3.37253